PROJECT ADMINISTRATION DATA SHEET

[X] ORIGINAL [ ] REVISION NO. ______

Project No. M-50-605 GTRI/~~GIT~~ DATE 7 /25 /83

Project Director: Dr. M. J. Sobel School/~~Lab~~ Management

Sponsor: National Science Foundation

Type Agreement: Grant No. ECS-8305963

Award Period: From 8/1/83 To 1/31/85* (Performance) 4/31/85 (Reports)

Sponsor Amount:

	This Change	Total to Date
Estimated:	$ 59,999	$ 59,999
Funded:	$ 59,999	$ 59,999

Cost Sharing Amount: $ 2,583 Cost Sharing No: M-50-322

Title: "Variability and Markov Decision Processes"

ADMINISTRATIVE DATA

OCA Contact John Burdette X4820

1) Sponsor Technical Contact:

Elias Schutzman, Acting Director
Dir. for Systems Theory & Operations Res.
Division of Electrial Computer & Systems
Engineering-National Science Foundation
1800 G Street, N.W.
Washington, D.C. 20550 (202) 357-9618

2) Sponsor Admin/Contractual Matters:

Mary Frances O'Connell
National Science Foundation
1800 G Street, N.W.
Washington, D.C. 20550

(202) 357-9602

Defense Priority Rating: N/A Military Security Classification: N/A

(or) Company/Industrial Proprietary: ______

RESTRICTIONS

See Attached NSF Supplemental Information Sheet for Additional Requirements.

Travel: Foreign travel must have prior approval — Contact OCA in each case. Domestic travel requires sponsor approval where total will exceed greater of $500 or 125% of approved proposal budget category.

Equipment: Title vests with GIT

COMMENTS:

*Includes usual 6 month unfunded flexibility period.

COPIES TO:

Project Director
Research Administrative Network
Research Property Management
Accounting

Procurement/EES Supply Services
Research Security Services
Reports Coordinator (OCA)
Research Communications (2)

GTRI
Library
Project File
Other I. Newton

FORM OCA 4:383

GEORGIA INSTITUTE OF TECHNOLOGY OFFICE OF CONTRACT ADMINISTRATION

SPONSORED PROJECT TERMINATION/CLOSEOUT SHEET

Date 4/18/88

Project No. M-50-605 School/Lab MGMT

Includes Subproject No.(s) N/A

Project Director(s) M. J. Sobel GTRC/~~GIT~~

Sponsor National Science Foundation

Title "Variability and Markov Decision Processes"

Effective Completion Date: 1/31/86 (Performance) 4/30/86 (Reports)

Grant/Contract Closeout Actions Remaining:

- [ ] None
- [ ] Final Invoice or Copy of Last Invoice Serving as Final
- [ ] Release and Assignment
- [X] Final Report of Inventions and/or Subcontract: Patent and Subcontract Questionnaire sent to Project Director [X]
- [ ] Govt. Property Inventory & Related Certificate
- [ ] Classified Material Certificate
- [ ] Other

Continues Project No. Continued by Project No.

COPIES TO:

Project Director
Research Administrative Network
Research Property Management
Accounting
Procurement/GTRI Supply Services
Research Security Services
Reports Coordinator (OCA)
Program Administration Division
Contract Support Division
Facilities Management - ERB
Library
GTRC
Project File
Other

Progress Report on Grant No. ECS-8305963,

"Variability and Markov Decision Processes"

I. ORIGINAL OBJECTIVES

A Markov decision process (abbreviated MDP) is a controlled stochastic process. The usual MDP optimization criteria are the expected values of random variables such as present value, total reward, and the typical period's reward in the long-run. Other characteristics of the distribution functions of these random variables, such as the variance, are important in engineering practice. Also, intertemporal variation is important in many phenomena modeled with MDP's. The proposed research seeks to develop efficient methods to compute the distribution functions of the criteria random variables, to characterize intertemporal variability, and to delineate tradeoffs between expected values and other characteristics such as variances. Attention will be paid to MDP models in applications such as water resource and water quality systems, inventory and production processes, and coastal fisheries.

The original research proposal provided more detailed objectives in six subareas. This progress report will use the same organization as part II of the original proposal.

II. RESEARCH OBJECTIVES, PROGRESS, AND PLANS

A. Mean-Variance Tradeoffs

The goal here is to develop algorithms to specify mean versus variance tradeoffs in MDP's. Preliminary results have been obtained and the PI plans to write a paper on this topic during the forthcoming year.

B. RISK-RELATED CRITERIA

The goal here is to study the maximization of the expected utility of the present value of rewards in an MDP. That is, let B_s indicate the present value of rewards starting from initial state s. Then the goal is to study the maximization of $E[u(B_s)]$ for appropriate functions u(•). The proposal identified two important problems:

(i) For what class of functions u(•) does there exist a stationary policy which is maximal (for all initial states s)?

(ii) How can an optimal policy be computed efficiently?

In "Nonstationary Policies are Optimal for Risk-Sensitive Markov Decision Processes," completed during the past year by the PI and M. Bouakiz, a Ph.D. student doing research under the grant, the following fundamental result has been obtained. If u(•) is nonlinear then there is a finite-state MDP for which a nonstationary policy strictly dominates all stationary policies (or randomizations of stationary policies). "Nonlinear" means that u(•) has a second-order Taylor approximation at a point where the second-derivative (in the approximation) is nonzero. Thus, the answer to (i) above is that only linear functions u(•) necessarily have stationary optimal policies. The manuscript cited above also shows that stationary policies can be strictly suboptimal if u(•) is nonlinear, the discount factor is unity, and the MDP is terminating.

In "Sensitive Risk-Sensitive MDP's," a draft paper completed under the grant, the PI shows that the discouraging answer to question (i) above does not prevent the investigation of computational issues. Specifically, various formulas are derived for post-optimality analysis of risk-neutral models. The question is: how will the introduction of a "little" risk sensitivity alter the optimality of already computed policies. It is possible to parameterize a "little" risk sensitivity in a manner that would apply to a broad class of utility functions. The formulas identify the manner in which the risk-neutral optimal policy would change as the risk parameter grows larger.

C. SPECIAL STRUCTURES AND MOMENT FORMULAS

The goal here is to exploit the structure of classes of models and simple policies (e.g., inventory models with (s,S) optimal policies) in order to obtain analytical expressions for the distribution function of the present value of the time-stream of rewards. Also, the goal is to obtain formulas for the moments of the distribution functions. Research on these topics is at an early stage. Several models in quality control and in production control are in early stages of investigation by the PI and K-J. Chung, a Ph.D. student whose research has been partly supported by the grant.

D. EXPANSIONS AND APPROXIMATIONS

Three analytical approximation approaches are described (in the proposal) for the goal of approximating the distribution function of the present value of the time-stream of rewards. Only one of these, namely the truncated moment problem of Chebyshev and Markov, has been investigated. "Linear Programming Solutions of Truncated Moment Problems," a manuscript completed by the PI and K-J. Chung this past year, concern the classical truncated moment problem for a scalar random variable. It is shown that if the random variable's support set is finite, then the classical problem is a linear programming problem of relatively small size. Then the paper presents the results of a simulation study of the effectiveness of the truncated moment problem and related heuristics for estimating an unknown distribution function. The results seem important and will be presented this summer at the national meeting of the American Statistical Association.

The truncated moment problem for a Markov chain with rewards, described in the research proposal, is really a vector rather than a scalar problem. During the forthcoming year the PI plans to apply the lessons learned on the scalar problem to the vector problem. Then the plan is to contrast the resulting methods with simulation methods developed by Professor George S. Fishman at the University of North Carolina at Chapel Hill. The PI plans to develop a detailed experimental plan for this phase of the project during the summer of 1984.

E. ENGINEERING APPLICATIONS

The goal of this research is to apply the general methodological results to models of engineering processes. M. Bouakiz, a Ph.D. student whose research is supported by the grant, has written drafts of three papers on engineering processes. One, on equipment replacement models, considers the extent to which binary decision MDP's with exponential utility functions have first-period decision rules which have the "control-limit" property. His work on inventory models investigates the monotonicity of optimal ordering rules. His research on reservoir operating rules seeks conditions under which the optimal release rules would have a simple form.

K-J. Chung and the PI have begun to investigate several models concerned with quality control and with production scheduling. The research has not advanced far enough to report on the results.

F. OTHER APPLICATIONS

The proposal observed that risk sensitivity is apparent in the design and operation of many non-engineering processes. M. Bouakiz has preliminary results concerning a capital accumulation model of a very general form. The PI has written a draft of "Optimal Barrier Dividend Policies in One-Person Survival GAmes." The underlying model concerns insurance and finance operations in which a risk of bankruptcy is present. The manuscript shows that "pathological" counterexamples obtained by other authors are explainable as part of a general theory.

FINAL TECHNICAL REPORT

DUE 4/30/86

DELIVERABLE

3

StonyBrook

W. Averell Harriman School
for Management and Policy
State University of New York at Stony Brook
Stony Brook, New York 11794-3775
telephone: (516) 632-7180

M-50-605 / Sobel
Management.

Office of Contract Administration
Georgia Institute of Technology
225 North Avenue
Atlanta, GA 30332

To whom it may concern:

RE: Overdue deliverables report, 12/31/87

The report (copy enclosed) lists overdue deliverables in connection with M-50-605 and M-50-607, projects for which I was the PI/PD.

First, concerning M-50-605, it is more than a year since a final report was submitted. As you know, NSF requests a section on "Results from Prior NSF Support" when a follow-on proposal is submitted and, in such a situation, that section constitutes the final report on the expiring grant. I have enclosed a copy of the ten-page section on "Results from Prior NSF Support" which was included in a new proposal which was sumitted more than a year ago. The new grant was funded and its NSF number is ECS-8705698. The Director of the Program which awarded the new grant is Dr. Marlin U. Thomas and his phone number is 202-357-9618.

Second, concerning M-50-607, this grant was in effect jointly awarded to Martin Shubik of Yale University and myself. He is about to submit a follow-on proposal to NSF which will contain a report on the expired grant activities both at Georgia Tech and Yale. I will send you a copy of the section of his proposal which contains the report shortly.

Sincerely,

Matthew J. Sobel

cc: Dean Gerald J. Day, College of Management

XC: RAN (Cooper)
Welch
Heitmuller (2 copies)
file

for sequential games; generally their prospects for computational solutions are far poorer than for MDPs [cf. Parthesarathy and Raghaven (1981)]. Examples of sequential games whose solutions have been extracted via the myopic approach include [S16, S24, S30, S43].

The numerical solution of particular models is another attraction of models with myopic optima. It will often be feasible to optimize the multidimensional static model although "the curse of dimensionality" would have precluded direct numerical solution of the dynamic model. This approach to exorcising the curse has been exploited in several types of dynamic optimization problems including inventory, maintenance, advertising, capacity expansion, and transportation networks.

II. RESULTS FROM PRIOR NSF SUPPORT

NSF Grant ECS-8305963 was awarded in 1982 and expired December 31, 1985.[3] Its topic was "Variability and Markov Decision Processes" and a portion of the research proposed here builds on the directions pursued in the previous grant. The connections are most apparent in parts IA and IIIA of this proposal. Other portions are not closely related to the previous grant. This section of the proposal reports on the results of the earlier grant.

---

[3] This section of the proposal does not report on results from NSF Grant SES-3408951 on which Sobel is co-principal investigator with Martin Shubik (Yale University). Not only is the grant administered by a different NSF program, but the topic is largely unrelated either to the current proposal or ECS-8305963. The grant will expire January 31, 1987 and concerns models of computer-based business games and related oligopoly models.

A. Research Objectives in Prior Support

The research objectives of the prior grant were listed in part II of the proposal for that grant and they will be summarized here.

1. Mean-Variance Tradeoffs

The principal goal here was to develop algorithms to specify mean versus variance tradeoffs in the sense of Pareto optimality for average-reward criterion MDPs. This bicriterion problem differed from those analyzed earlier by Henig (1985a), White and Kim (1980), and others. Two closely related problems which were mentioned were (a) is there a loss of Pareto optimality in confining comparisons to randomized stationary policies, and (b) can algorithms for mean-variance tradeoffs be developed for discounted-return MDPs?

2. Risk-Related Criteria

Recall the notation B_s for the random variable which is the present value of the time stream of rewards in an MDP whose initial state is $s_1 = s$ (and with $K = L = 1$). If a decision maker is sensitive both to the timing of rewards and to risk, then there are axiomatic justifications for considering the objective of maximizing $E[u(B)]$ where the nonlinear utility function $u(\cdot)$ reflects the decision maker's attitude towards risk. The primary goals here were (a) to identify interesting classes of functions $u(\cdot)$ for which a stationary policy is maximal for the optimization problem, and (b) to find algorithms for computing optimal policies. The literature in decision theory, economics, and operations research [e.g., Howard and Matheson (1972), Jaquette (1976), Porteus (1975), and Denardo and Rothblum (1977)] suggest that particular attention be paid to exponential and quadratic utility functions.

3. Special Structures and Moment Formulas

Recall that F_s is the distribution function of the present value B_s in an MDP with K = L = 1 if a stationary policy is employed. The functional equation (2) for $\{F_s : s \in \Omega\}$ has formidable obstacles to easy analysis. The objective here was to determine if special structures in the transition probabilities of the MDP might simplify the computation of the distribution functions or their moments.

4. Expansions and Approximations

The functional equation (2) mentioned in the previous paragraph can be exploited to obtain a recursive matrix equation for the seqence $\underset{\sim}{v}^{(1)}$, $\underset{\sim}{v}^{(2)}$, ... of vectors of moments of the present value of an MDP (with K = L = 1) if a stationary policy is employed. This recursion suggests that it might be more efficient to compute a partial sequence of moments and to utilize the partial sequence to approximate the distribution functions $\{F_s : s \in \Omega\}$ rather than to compute (approximate) the distribution functions directly. This possibility is closely related to (a) the classical truncated moment problem of Markov and Chebyshev [cf. Shohat and Tamarkin (1943)] and (b) more recent multivariate versions of the moment problem [e.g., Godwin (1964), Marshall and Olkin (1979), and Tonge (1980), and page 323 in Volume 4 of Johnson and Kotz (1969-1972)]. The principal goal here was to investigate the relative efficacy of moment approximations, contraction mapping approaches, and simulation avenues to approximate the distribution functions of the present values.

5. Engineering Applications

The objective here was to apply the general results obtained under headings #1 through #4 to some structured models which arise in various applications. The possible applications mentioned were reservoir management, stormwater retention basins, production planning, multiserver queueing processes, vehicular traffic, and communications networks.

6. Other Applications

MDPs have been used to model many non-engineering phenomena. The goal here was to apply the general results under headings #1 through #4 to structured models of some non-engineering phenomena. The possible applications mentioned were capital accumulation, fisheries management, and financial portfolio selection.

B. List of Publications

This section lists the research reports which were written under the sponsorship of the previous grant.

1. "Discounted MDP's: Distribution Functions and Exponential Utility Maximization," K-J. Chung and M.J. Sobel, SIAM J. Opt. & Appl., to appear.

2. "Mean-Variance Tradeoffs in an Undiscounted MDP," M. J. Sobel, Operations Research, under revision.

3. "Maximal Mean/Standard Deviation Ratio in an Undiscounted MDP," M. J. Sobel, Operations Research Letters, Vol. 4 (1985), 157-159.

4. "Accelerating Mean-Variance Tradeoffs in an Undiscounted MDP," K-J. Chung, Operations Research, in process.

5. "The Target-Level Criterion in MDPs," M. Bouakiz, JOTA, in process.

6. "Nonstationary Policies are Optimal for Risk-Sensitive MDPs," M. Bouakiz and M.J. Sobel, Math. of Oper. Res., in process.

7. "Replacement Policies under Risk-Sensitive Criteria," M. Bouakiz, Oper. Res. Quart., in process.

8. "Inventory Control with an Exponential Utility Criterion," M. Bouakiz and M. J. Sobel, Operations Research, in process.

9. "Production Variability vs. Cost," M. J. Sobel.

10. "Linear Programming Solutions of the Truncated Moment Problem," K-J. Chung and M. J. Sobel.

C. Summary of Results

1. The first half of "Discounted MDPs: Distribution Functions and Exponential Utility Maximization" applies fixed-point theory to (2), a functional equation for the distribution functions of the present value. Suppose H is a finite set (an assumption made here for expository convenience), K = L = 1, and a stationary policy δ^∞ is used. Let Y be the set of distribution functions on the smallest (finite) interval U which contains the support sets of F_s for all $s\epsilon\Omega$. Let $Z = \times_{s\epsilon\Omega} Y$. Then the vector $F_s = (F: s\epsilon\Omega)$ (the ordering in Ω is arbitrary) is an element of Z. Let V* be the set of all continuous linear functionals on the space of all real-valued continuous functions on U, and let $W = \times_{s\epsilon\Omega} V^*$. A mapping M: $Z \to W$ can be defined via (2) as follows for $s\epsilon\Omega$ if $G = [G_i(\cdot);\ i\epsilon\Omega]$:

$$M(G)_s(x) = \Sigma_{(j,k)\epsilon H}\, q_{sjk}\, G_j[(x-k)/\beta] \tag{4}$$

It can be shown that M maps Z into itself and G is a fixed point of M if, and only if, G satisfies (2) for all $s\epsilon\Omega$.

The first half of this paper shows that M is a nonexpansive, weakly continuous, and affine mapping but that it is not necessarily a contraction mapping. Therefore, successive approximations will not necessarily solve (2). An example is given in which (4) has multiple fixed points! On the other hand, (a) an initial point in Z is specified such that successive approximations converges monotonely to F, and (b) sufficient conditions are given for an easily computed sequence of elements of Z to have a subsequence which converges to F. This part of the paper responds to objectives previously listed in part A4.

The second half of this paper concerns maximization of E[u(B)] for the exponential utility function $u(x) = -\exp(-\lambda x)$ ($\lambda > 0$) which has constant local risk aversion and constant risk premium. It is shown that this optimization problem corresponds to the dynamic programming-type functional equation

(5) $$f(s,\lambda) = \max\{\Sigma_{(j,k)\epsilon H}\, m^a_{sjk} f(j,\beta\lambda) : a\epsilon A_s\} \qquad (s\epsilon\Omega, \lambda>0)$$

where $m^a_{sjk} = \exp(-\lambda k)q^a_{sjk}$. It is shown that the value function can be approximated monotonely by successive approximations starting from an easily computed initial approximation. A family of bounds on the value function is specified which can be employed to decide when to truncate the iterative procedure to approximate the value function. This part of of the paper responds to previous objectives listed in part IIA2.

2. The problem addressed in "Mean-Variance Tradeoffs in an Undiscounted MDP" is the computation of Pareto optima, in the sense of high mean and low variance of the stationary distribution, in the stationary probability distribution of the reward in an MDP (with K = L = 1) controlled by a stationary policy. The paper completely solves the unichain case; parametric analysis of a linear program having the same number of variables and one more constraint than the familiar formulation for gain-rate optimization yields the solution. The paper presents partial results for the same linear program in the multichain case if the initial state is an element of choice. Therefore, most of the computational labor of Pareto optimization has already been expended when the gain-rate optimization problem has been solved. The paper responds to previous objectives in part IIA1 of the proposal.

3. "Maximal Mean/Standard Deviation Ratio in an Undiscounted MDP" concerns maximization of the mean/standard deviation ratio of the stationary distribution of the reward mentioned in the previous paragraph. Parametric analysis of essentially the same linear program mentioned in the previous paragraph is the principal element of an algorithm which yields a maximum value of the ratio. This problem had been treated erroneously previously; see Miller (1978) for comments on the errors. The paper responds to (previous) objectives in part IIA1.

4. "Accelerating Mean-Variance Tradeoffs in an Undiscounted MDP" shows that the range of parametric analysis in paper #2 above can be reduced without loss of Pareto optimality. Moreover, the starting point of the pertinent range can be computed easily. The paper's method of analysis is to derive properties of the extreme points of the gain-rate optimization linear program which are pertinent to mean-variance tradeoffs. The method of this paper can be applied to the algorithm in paper #3 to accelerate it too. The paper responds to objectives in part A1.

5. "The Target-Level Criterion in MDPs" concerns maximization of E[u(B)] when u(·) is a step function, i.e., maximization of the probability that the present value of the rewards is at least as great as a target level. This criterion has been employed in economics where it is called "satisficing," in gambling, and in reliability theory where it is central to the definition of "reliability." The paper shows that optimal policies are usually nonstationary (this result is <u>not</u> encompassed by paper #6 below) and derives functional and recursive approximations which correspond to optimal policies. It is shown that successive approximations can be employed to approximate an optimal policy if an easily computed initial approximation is employed. Also, bounds on the value function of the functional equation are specified which would facilitate the truncation of computations when a nearly optimal policy has been computed. This paper responds to (previous) objectives described in part A2.

6. "Nonstationary Policies are Optimal for Risk-Sensitive MDPs" addresses the question of the existence of stationary optimal policies for infinite-horizon discounted MDPs with risk-sensitive criteria. The optimization problem is to maximize E[u(B)] where u(·) is nonlinear. J. M. Harrison [see Jaquette (1976)] had shown that a nonstationary policy could dominate all stationary policies when u(·) is exponential and [S31]

provided grounds to expect a similar result if $u(\cdot)$ is quadratic. This paper considers smooth nonlinear utility functions, i.e., suppose there exists $M > 0$ and x_o such that

$$u(x) = u(x_o) + u'(x_o)(x-x_o) + (x-x_o)^2 u''(x_o)/2 + o[(x-x_o)^2]$$

for all $|x-x_o| < M$ and $u''(x_o) \neq 0$. Under this assumption, one can construct an MDP with finitely many states and actions for which a nonstationary policy is strictly better than any stationary policy (or randomized stationary policy). This result raises research issues which are discussed in Section IIIA of this proposal.

7. "Replacement Policies under Risk-Sensitive Criteria" examines the structure and computation of optimal policies in replacement models with exponential or target-level (see the discussion of #5 above) utility functions. The MDPs studied here possess control-limit policies which are optimal for the risk-neutral criterion, i.e., maximize E(B). Necessary and sufficient conditions for the exponential and target-level criteria are given which ensure that a nonstationary control-limit policy is optimal. The paper also presents results concerning the sensitivity of the control limits with respect to changes in the risk parameters (exponential parameter target level).

8. "Inventory Control under Risk-Sensitive Criteria" studies single-product dynamic inventory models with stochastic demand which possess optimal base-stock policies for a risk-neutral criterion. It is shown that an exponential utility function generally preserves the qualitative structure of the optimal policy, the optimal base-stock levels are nonstationary, and each period's optimal base-stock level can be computed with a static optimization problem. A target-level criterion is more difficult to analyze. The single-period problem is completely analyzed but only partial results are presented for the multiperiod model.

Even with linear holding and penalty costs and exponentially distributed demand, the value functions associated with appropriate recursive equations will generally be neither convex nor concave functions of the inventory level at the beginning of the period.

9. "Production Variability vs. Cost" is motivated by the difficulty of estimating smoothing cost parameters in applications of production smoothing models. The usual response to this difficulty is to perform extensive sensitivity analysis with respect to the smoothing cost parameters. The approach taken in this paper is to employ the long-run variance of the production quantity as an alternative to monetary smoothing costs. It is shown that the tradeoff between this variance and the long-run average cost per period (not including smoothing costs) in a backorder production model can be explored via parametric analysis of a linear program. The linear program is relatively small and Pareto-optimal policies possess monotonicity properties which can be exploited to accelerate computations.

10. "Linear Programming Solutions of the Truncated Moment Problem" is a first-step in the examination of the approximations issues summarized in part IIA4 of this proposal. The goal is to approximate the distribution functions of the present values, F_s $(s \epsilon \Omega)$, by viewing the problem as an extension of multivariate versions of the truncated moment problem to a Markov chain version of the same problem. However, the literature on the univariate and (multivariate) problem is surprisingly devoid of algorithms. This paper presents a small linear program for the univariate truncated moment problem with a finite support set. Since truncated moment problems are typically prompted by the desire to estimate an unknown distribution function, the paper describes the results of a simulation experiment which tests the effectiveness of various estimators and some related heuristics.

D. Relation of Completed Work to Proposed Work

Approximately half of the proposed work is concerned with risk sensitivity and, in that sense, is related to the completed work. The objectives of the related half are presented in part IIIA of this proposal. It will be apparent that most of these objectives are fundamentally new tasks rather than "tidying up" and extending the completed work. The connections between the completed work and the proposed work are explained here briefly and are discussed in greater detail in part IIIA.

One of the two principal objectives in part IIIA1 concerns clarification of the multichain case in the paper [S36] which was summarized in Part IIC2 The other objective in IIIA1 is motivated by the prospect of exploiting IIC2, IIC3, and IIC4 but the methods and direction are quite different from those in the completed work.

Intertemporal variation is balanced against various costs in the production model considered in [S40] which was summarized in part IIC9. However, most of the objectives in IIIA2 concern problems and models which are not found in [S40].

Part IIC6 summarizes the use of a Taylor-series approximation for the utility function in [S35]. However, the objectives in part IIIA3 consist of different kinds of expansions which would be utilized for purposes unrelated to those in [S35].

The objectives in part IIIA4 concern risk sensitivity in sequential games. The completed work is not related to these objectives except to the extent that it suggests parallel issues which merit investigation.

Some of the objectives in part IIIA5 are closely related to the completed work. The completed work summarized in parts IIC5 and IIC6 shows that nonlinear utility functions which are either smooth or step functions generally induce nonstationary optimal policies. A minor objective suggested by these results is to devise a unified proof for these two results. A more important objective is to obtain bounds on the quality of stationary policies as approximations to optimal nonstationary policies. Another objective in part IIIA5 is related to the completed work summarized in parts IIC7 and IIC8 which concerned risk sensitivity in replacement and inventory models, respectively. The objective is to unify, extend, and generalize the results on the monotonicity of parameters of optimal nonstationary policies as functions of the degree of risk sensitivity.

. The objectives in part IIIB are not related to the completed work.

III. RESEARCH OBJECTIVES

This section of the proposal describes the objectives of the proposed research. The parts concern risk sensitivity (IIIA) and myopic-affine models (IIIB).

A. Risk Sensitivity

1. Mean-Variance Tradeoffs

Recall the notation v_s and V_s for the mean and variance, respectively, of the present value B_s if the initial state is s (and K = L = 1). A form of Pareto optimality concerning undiscounted counterparts of v_s and V_s was investigated in [S36]. The strongest results in that paper concerned the unichain case and decision variables in the style of Derman (1962), namely elements (x_{sa}) of the set X where $x_{sa} \geq 0$, $\Sigma_s \Sigma_a \; x_{sa} = 1$, and $\Sigma_s \Sigma_a \; (\delta_{sj} - p^a_{sj}) x_{sa} = 0$ for all j (where $\delta_{sj} = 1$ if $s = j$ and $\delta_{sj} = 0$ if $s \neq j$). The correspondence between X and the set of stationary policies is clear in the unichain case; so the results in [S36] have a clear policy-space interpretation in the unichain case. In the multichain case (which